

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 841 781 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
13.05.1998 Bulletin 1998/20

(51) Int Cl. 6: H04L 12/44

(21) Application number: 97630053.3

(22) Date of filing: 05.09.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

Designated Extension States:
AL LT LV RO SI

(30) Priority: 11.09.1996 US 714029

(71) Applicant: McData Corporation
Broomfield, Colorado 80021-3464 (US)

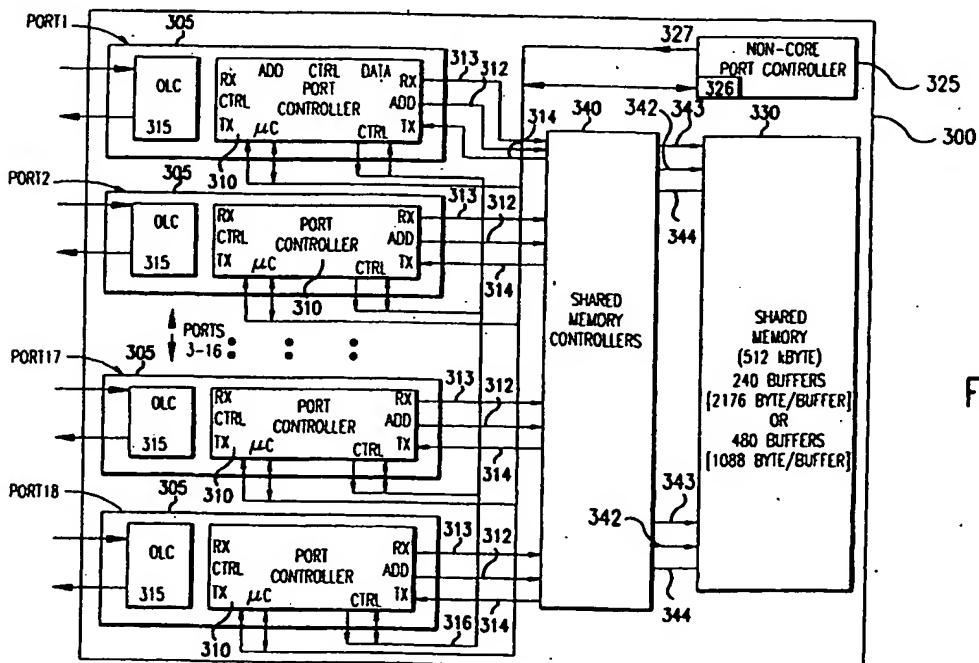
(72) Inventor: Book, David
Thornhill, Ontario L4J 7X1 (CA)

(74) Representative: Waxweiler, Jean et al
Dennemeyer & Associates S.A.,
P.O. Box 1502
1015 Luxembourg (LU)

(54) Fibre channel switch employing distributed queuing

(57) The present invention is a fibre channel switch employing a distributed queuing algorithm for interconnecting a plurality of devices (workstations, supercomputer, peripherals) through their associated node ports (N_ports) and employs a fabric having a shared memory coupled to a plurality of fabric ports (F_ports) through a bi-directional bus over which memory addresses, frame data and communications commands are transmitted.

Each F_port includes a port controller employing a distributed queuing algorithm associated with a control network for communicating commands between the ports related to when and where frame transfers should be made, wherein the bi-directional bus provides an independent data network for access to the shared memory such that frames can be transferred to and from the shared memory in response to port controller commands.



Description**FIELD OF THE INVENTION**

The present invention relates an apparatus for distributed source and destination queuing in a high performance memory based switch and, more particularly, to a Fibre channel switch having a distributed queuing algorithm.

BACKGROUND OF THE INVENTION

Mainframes, super computers, mass storage systems, workstations and very high resolution display subsystems are frequently connected together to facilitate file and print sharing. Common networks and channels used for these types of connections oftentimes introduce communications bottle necking, especially in cases where the data is in a large file format typical of graphically-based applications.

There are two basic types of data communications connections between processors, and between a processor and peripherals. A "channel" provides a direct or switched point-to-point connection between communicating devices. The channel's primary task is merely to transport data at the highest possible data rate with the least amount of delay. Channels typically perform simple error correction in hardware. A "network," by contrast, is an aggregation of distributed nodes (e.g., Workstations, mass storage units) with its own protocol that supports interaction among these nodes. Typically, each node contends for the transmission medium, and each node must be capable of recognizing error conditions on the network and must provide the error management required to recover from the error conditions.

One type of communications interconnect that has been developed is Fibre Channel. The Fibre channel protocol was developed and adopted as the American National Standard for Information Systems (ANSI). See *Fibre Channel Physical and Signaling Interface*, Revision 4.2, American National standard for Information Systems (ANSI) (1993) for a detailed discussion of the fibre channel standard. Briefly, fibre channel is a switched protocol that allows concurrent communication among workstations, super computers and various peripherals. The total network bandwidth provided by fibre channel is on the order of a terabit per second. Fibre channel is capable of transmitting frames at rates exceeding 1 gigabit per second in both directions simultaneously. It is also able to transport commands and data according to existing protocols such as Internet protocol (IP), small computer system interface (SCSI), high performance parallel interface (HIPPI) and intelligent peripheral interface (IPI) over both optical fiber and copper cable.

Fig. 1 illustrates a variable-length frame 11 as described by the Fibre Channel standard. The variable-length frame 11 comprises a 4-byte start-of-frame (SOF)

indicator 12, which is a particular binary sequence indicative of the beginning of the frame 11. The SOF indicator 12 is followed by a 24-byte header 14, which generally specifies, among other things, the frame source address and destination address as well as whether the frame 11 is either control information or actual data. The header 14 is followed by a field of variable-length data 16. The length of the data 16 is 0 to 2112 bytes. The data 16 is followed successively by a 4-byte CRC (cyclic redundancy check) code 17 for error detection, and by a 4 byte end-of-frame (EOF) indicator 18. The frame 11 of Fig. 1 is much more flexible than a fixed frame and provides for higher performance by accommodating the specific needs of specific applications.

FIG. 2 illustrates a block diagram of a representative fibre channel architecture in a fibre channel network 100. A workstation 120, a mainframe 122 and a super computer 124 are interconnected with various subsystems (e.g., a tape subsystem 126, a disk subsystem 128, and a display subsystem 130) via a fibre channel fabric 110 (i.e. fibre channel switch). The fabric 110 is an entity that interconnects various node-ports (N_ports) 140 and their associated workstations, mainframes and peripherals attached to the fabric 110 through the F_ports 142. The essential function of the fabric 110 is to receive frames of data from a source N_port and, using a first protocol, route the frames to a destination N_port. In a preferred embodiment, the first protocol is the fibre channel protocol. Other protocols, such as the asynchronous transfer code (ATM) could be used without departing from the scope of the present invention.

Essentially, the fibre channel is a channel-network hybrid, containing enough network features to provide the needed connectivity, distance and protocol multiplexing, and enough channel features to retain simplicity, repeatable performance and reliable delivery. Fibre channel allows for an active, intelligent interconnection scheme, known as a "fabric," or fibre channel switch to connect devices. The fabric includes a plurality of fabric-ports (F_ports) that provide for interconnection and frame transfer between a plurality of node-ports (N_ports) attached to associated devices that may include workstations, super computers and/or peripherals. The fabric has the capability of routing frames based upon information contained within the frames. The N_port manages the simple point-to-point connection between itself and the fabric. The type of N_port and associated device dictates the rate that the N_port transmits and receives data to and from the fabric. Transmission is isolated from the control protocol so that different topologies (e.g., point-to-point links, rings, multidrop buses, cross point switches) can be implemented.

The Fibre Channel industry standard also provides for several different types of data transfers. A class 1 transfer requires circuit switching, i.e., a reserved data path through the network switch, and generally involves

the transfer of more than one frame, oftentimes numerous frames, between two identified network elements. In contrast, a class 2 transfer requires allocation of a path through the network switch for each transfer of a single frame from one network element to another.

Frame switching for class 2 transfers is more difficult to implement than class 1 circuit switching as frame switching requires a memory mechanism for temporarily storing incoming frames in a source queue prior to their routing to a destination port, or a destination queue at a destination port. A memory mechanism typically includes numerous input/output (I/O) connections with associated support circuitry and queuing logic. Additional complexity and hardware is required when channels carrying data at different bit rates are to be interfaced.

It is known to employ centralized queuing that is inherently slow as a common block of logic must be employed for all routing decisions within the switch.

It is also known to employ distributed source queuing which has apparent disadvantages when the frame at the head of the queue is destined to a port that is already forwarding a frame such that the path is blocked and the frame cannot be transferred. Alternatively, it is known to employ distributed destination queuing, which has the apparent disadvantage of a large destination queue at each port as it is possible for all frames within the switch to be simultaneously queued to the same destination port.

Another disadvantage of distributed destination queuing is apparent when the frame at the end of the head of the queue is sourced from a port that is already forwarding a frame such that the path is blocked and the frame cannot be transferred.

Thus, a heretofore unaddressed need exists in the industry for new and improved systems for implementing the Fibre Channel industry standard for class 2 transfers on fiber optic networks with much higher performance and flexibility than presently existing systems. Particularly, there is a significant need for a method and apparatus that combines both distributed source and destination queuing in a high performance memory based switch.

A need also exists to implement distributed queues between the source and destination ports, requiring the lower queue storage resources of source queuing, but providing the high throughput of destination queuing and avoiding "head-of-line" blocking of either source or destination queuing.

It would be desirable and of considerable advantage to provide a Fibre channel switch that provides for efficient transfer of queuing information between Fibre channel ports, especially if the new switch provided an improvement in any of the following areas: increased bandwidth, decreased no-load latency and increased throughput under load (due to parallelism of distributed queuing).

It will be apparent from the foregoing that there is still a need for a High-Bandwidth memory-based switch

employing distributed queuing that differs from that employed in existing centralized Fibre Channel switch architectures.

5 SUMMARY OF THE INVENTION

In accordance with the invention, there is provided a fibre channel switch employing distributed source and destination queuing for interconnecting a plurality of devices (workstations, supercomputer, peripherals) through their associated node ports (N_ports). The fibre channel switch provides a fabric having a shared memory coupled to a plurality of fabric ports (F_ports) through a bit-slicing memory controller over which memory addresses, frame data and communications commands are transmitted. Each F_port includes a port controller employing a distributed queuing algorithm associated with a control network for communicating commands between the ports related to when and where frame transfers are to be made. The bit-slicing memory controller provides an independent data network for access to the shared memory such that frames can be transferred to and from the shared memory in response to port controller commands.

Each port controller further comprises a link controller (LC) for controlling the optical interface with the N_port, a frame transfer controller (FTC) for receiving and transmitting data between the LC and the shared memory, an address validation /translation controller (AVTC) for obtaining frame header information and for generating destination port numbers, buffer numbers, and timeout values, and a writer and reader that together implement distributed source/destination queuing, a synchronizer (sync) for synchronizing port state transitions and the activities of the reader and writer, a statistics controller (SC) for counting statistics, and a port communications controller (PCC) for enabling port controllers to communicate.

Other aspects and advantages of the present invention will be come apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating by way of example the principles of the invention.

45 BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic representation of a variable-length frame communicated through a fiber optic switch of a fiber optic network in accordance with the Fibre Channel industry standard;

FIG. 2 depicts a block diagram of a representative Fibre Channel architecture.

FIG. 3 illustrates a block diagram of a Fibre Channel switch according to the present invention.

FIG. 4 illustrates a block diagram of a port controller located within the Fibre Channel switch illustrated in FIG. 3.

FIG. 5 illustrates the data organization for a bit-

sliced memory in accordance with the invention as embodied in a simplified four port fibre channel switch.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Briefly, the present invention provides a Fibre Channel switch based upon a distributed queuing algorithm for interconnecting a plurality of devices through their associated node ports (N_ports). The fibre channel switch provides a fabric having a shared memory coupled to a plurality of fabric ports (F_ports) through a bit-slicing shared memory controller through which memory addresses, frame data are transmitted. The Fibre Channel switch supports both circuit and frame switched connections for multiple baud rate interfaces.

FIG. 3 illustrates a block diagram of an Fibre Channel switch 300 having eighteen ports 305 numbered Port 1 to Port 18 (only Port 1, Port 2, Port 17 and Port 18 are illustrated) coupled to a 512 kByte shared memory 330 through a bit-slicing memory controller 340. Each of the eighteen ports 305 include a port controller 310 having memory address (ADD) lines 312 for transmitting memory addresses and Rx lines 313 for transmitting frame data to the bit-slicing memory controller 340, and Tx lines 314 for receiving frame data from the bit-slicing memory controller 340, and an OLM/GLC media adapter 315 for interfacing to the Fibre channel.

A bit-sliced memory architecture is disclosed in U.S. Patent Application Serial No. 08/330,279, filed October 27, 1994 and entitled "Channel Module for a Fibre Optic Switch with Bit Sliced Memory Architecture for Data Frame Storage" by Dwayne Bennett and is hereby incorporated by reference. The memory architecture disclosed in 08/330,279 teaches only bit slicing of the Rx frame data. The preferred embodiment of the invention employs bit-slicing of both the Rx frame data and the Tx frame data. This allows the shared memory 330 to be used not only for data frame storage, but also for the multiplexing function delivered by the bit-slicing memory controller 340.

An independent communications command bus 315 provides interconnections between the eighteen port controllers 310 and is employed for sending communication commands related to the timing and destination of frame transfers. The switch 300 also includes a single non-core port controller 325 interconnected with each port controller 310 over a micro controller bus 327.

FIG. 4 illustrates a block diagram of a port controller 310 illustrated in FIG. 3, comprising a link controller (LC) 410, a frame transfer controller (FTC) 420, an address validation & translation controller (AVTC) 430, a writer 440 and reader 450, a synchronizer (sync) 460, a statistics controller (SC) 470, and a port communications controller (PCC) 480.

The link controller 410 further comprises a control processor for OLC control (lock-to-ref, loopback control,

etc.), elastic store, 8B/10B conversion, work synchronization, ordered set encoding/decoding, port state machine (offline to online protocol, link recovery handshaking, etc.), CRC generation and checking, and frame validation.

5 The FTC 420 is an interface for receiving and transmitting data between the LC 410 and the shared memory controller 340. Upon receipt of data from the LC 410, the FTC 420 forwards the memory address and the data 10 to the shared controller memory 340. The memory address is based on the available source buffers. The FTC 420 tells the writer 440 which destination port and receive buffer the frame has been put in via the address validation/translation controller (AVTC) 430, such that 15 the designated receive buffer will not be reused until the reader 450 tells it to reclaim the receive buffer. When transmitting data, the FTC 420 gets a message from the reader 450 (originating from the control network, 20 through the port communication controller 480) to retrieve a frame from the shared memory 330 and send it to the destination F_port via the LC 410 and the OLC 315. The FTC 420 determines the memory address from the source port and buffer number contained in the message from the reader 450. When the FTC 420 passes 25 the EOF to the LC 410, it signals the reader 450 to indicate that the transfer is complete.

30 The AVTC 430 gets frame header information of inbound frames from the FTC 420. The frame header information includes: SOF, frame header words 1 and 2, and buffer number. The AVTC 30 also generates the following outputs to the writer 440: a destination port number, buffer number, forward frame timeout value, and busy frame timeout value and a flag indicating if the frame would be forwarded or rejected/discard. The 35 writer 440 and reader 450 of each port controller in the switch together implement distributed source/destination queuing. Each source port maintains a separate queue for each destination port, and each queue contains only frames received by that particular source port, resulting in a separate queue for each source/destination port combination in the switch. The queues are distributed because each source port passes the buffer descriptor at the head of its destination queues to the destination port before the source port is actually ready to 40 forward the frame. Thus, the destination port knows about the next frame that each of the source ports wants to forward to it next.

45 When the destination port begins forwarding a frame from a particular source port, the destination port requests the next buffer descriptor in the source port queue. Depending on the latency of the request and response between the two port controllers and the length of the frame being transmitted, the destination port will get the next buffer descriptor before or at approximately 50 the same time transmission of the current frame completes. By overlapping these two operations, a very high rate of sustained bandwidth can be maintained between two ports sending frames back and forth to each other.

When the destination port has seen the EOF command go out, the destination port sends a message to the source port telling the source port to free up the source buffer and send a receiver ready (RRDY) command.

Source blocking does not occur in a shared memory architecture as the destination port is free to choose a buffer from any source port. The destination port is therefore free to use any forwarding priority scheme the destination port desires. A simple round-robin scheme is employed for fairness. Alternatively, a priority scheme may be employed that gives higher priority to a particular source port, or to pace the traffic from a source port that is flooding the destination port with frames (i.e., class 3 traffic that isn't paced by end-to-end flow control) by servicing it less often than the other ports, for example, based on traffic history. The source prioritizes frames to each destination port based on frame type. For example, link control frames may be given higher priority than data frames in a simple first-come first-serve scheme.

Upon passing a buffer descriptor to the destination port, the source port controller must remember the frame. In particular, a busy frame signal must be sent back to source N_Port, or the frame must be discarded if the frame cannot be forwarded out of the destination port within some fixed time. In the preferred embodiment of the invention, the source port has responsibility for timing the frame upon passing of the buffer descriptor. Alternatively, the destination port may have responsibility for timing the frame and will forward a request to the source port to return a busy frame signal when the timer times out. The source port may be link reset, requiring the contents of the buffer to be discarded. The source port therefore still needs to be able to stop the destination port from forwarding the frame, and the same race condition exists between the destination port starting to forward the frame and the source port telling the destination port not to forward. The source port retains responsibility for timing the frame as the source port must time the frame until the source port buffer descriptor makes it to the head of the queue and is passed to the destination port.

When a port goes offline, there is a period of time during which the switch hardware must reject frames on its own until the address validation tables are updated by software. Rather than requiring a source port controller to be aware of the port state of all the other ports, frames will continue to be requested by the port controller of the offline port, and it will respond to a buffer descriptor from the source with a reject message. This also applies to frames sent to a port that is in link recovery.

The amount of logic required in the FTC portion of the port controller is reduced as the source port is not required to generate busy and reject frame commands. Instead, the buffer descriptor of the frame to be busied or rejected is queued to a "non-core" port controller 325 (FIG. 3) having an imbedded processor 326. Software is employed to manipulate the frame header, and queuing the frame back out to the original source port. While

the original frame is queued to the "non-core" port controller 325, the source continues to time the frame. If the frame is not forwarded within a predetermined period of time, the source port makes a request to the "non-core"

5 port controller 325 not to forward the original frame to the processor. If the frame does make it to the processor, and the reject/busy frame is queued back to the source, the "non-core" port controller must time the frame and discard it if necessary. In the preferred embodiment, the
10 timer within the "non-core" port controller should take into account the time between when the frame was first queued to the "non-core" port controller 325 and when it was actually delivered to it. The buffer information passed to the "non-core" port controller 325 includes a
15 time stamp. The processor time stamps frames upon receipt such that a time out value can be calculated that will ensure that in the worst case, the frame is discarded within R_A_TOV.

A synchronizer (sync) 460 is employed or synchronizing port state transitions and the activities of the reader 450 and writer 440. For example, when a port goes into link recovery, the synchronizer 460 prevents the link controller 410 from taking the link active again until the writer has freed up all of the source buffers by either
20 discarding the frames or waiting for a destination to complete transmission of a frame. The synchronizer 460 also waits for the reader to recognize that the port is no longer active and aborts the current forward operation
25 or waits until it is finished. The synchronizer 460 does basically the same thing when the port goes offline or into link failure, but it must also wait for software to tell it that the port has been logged out of the fabric before allowing the link controller 410 to take the link back online.

30 The statistics controller (SC) 470 is located between the reader/writer and the LC/FTC such that it can count statistics, and pass information through, for example, when the writer 440 releases a buffer, the SC 470 tells the FTC 420 that the buffer space is available
35 and also tells the LC 410 to send a receiver ready (RRDY) signal. When the reader 450 decides a frame can be forwarded, the SC 470 is given the source port and buffer number information, allowing it to count frames transmitted on a per destination port basis, and
40 it also passes the information on to the frame transfer controller so it calculates the shared memory address of the frame data. The SC 470 connection to the FTC 420 and LC 410 would also allow Fibre channel class
45 one error and word count statistics to be counted if required.
50

The port communication controller (PCC) 480 allows port controllers 310 to communicate with one another. More specifically, writers 440 communicate with readers 450 and vice-versa, but readers 450 don't communicate with one another, nor do writers 440. Each port controller 310 has a single serial output line that is connected to all the ports (including itself so that an N_Port can send frames to itself). In an 18 port switch, each port

ing four ports and a 512kByte shared memory, those skilled in the art will recognize that modification and variations may be made such that the invention is equally applicable to much larger numbers of ports and memory.

Claims

1. A fibre channel switch for interconnecting a plurality of devices, each device having a node port (N_port), such that frames of data originating from a source port having a source N_port may be transferred through the fibre channel switch to a destination device having a destination N_port in accordance with a distributed source and destination queuing algorithm, comprising:

a plurality of fabric ports (F_ports), each F_port coupled to a different device through the corresponding source or destination N_port,
a plurality of port controllers, each port controller located on a different F_port, each port controller further comprising a distributed source and destination queuing algorithm for queuing frames sourced or destined from a device, or to, itself, and wherein, each port controller operates independently of the frame queuing by any other port controller,
a control network for communicating timing and destination commands between each port controller corresponding to when and where frame transfers are to be made,
a shared memory controller coupled to each port controller, for bit-slicing data and memory addresses being transmitted between the port controllers and the shared memory, wherein data and destination addresses are transferred to and from the shared memory,

receives 1 bit of data at a time to, and from, each memory bank, and over $2p$ clock cycles these bits are accumulated, and wherein, every F_port has access to the shared memory.

- 25
4. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 1, each port controller further comprising a writer and a reader, each writer further containing a separate destination queue for each destination port, wherein each source port passes the buffer descriptor at the head of its destination queues to the destination port before the source port is actually ready to forward the frame such that each destination port knows about the next frame that each of the source ports wants to forward next.
- 30
35
40
45
50
55
6. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 4, wherein upon forwarding a frame from a particular source port to a device, the destination port requests the next buffer descriptor in the source port queue, and wherein, depending on the latency of the request and response between the two port controllers and the length of the frame being transmitted, the destination port will get the next buffer descriptor before or at approximately the same time transmission of the current frame completes, and wherein, by overlapping these two operations, a very high rate of sustained bandwidth can be maintained between two ports sending frames back and forth to each other.
6. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 2, wherein $p=18$ F_ports and $b=800$ Mbits/sec, $n=36$ memory devices and $mb=800$ Mbits/sec, and wherein, the tx side of all

- ports retrieve data directly from the same memory which the rx side accesses.
7. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 1, each port controller further comprising:
- 5 a link controller (LC) for controlling the link between the N_port and the F_port,
 - 10 a frame transfer controller (FTC) for receiving and transmitting data between the LC and the shared memory,
 - 15 an address validation/translation controller (AVTC) for obtaining frame header information and for generating destination port numbers, buffer numbers, and timeout values,
 - 20 a synchronizer (sync) for synchronizing port state transitions and the activities of the reader and writer, a statistics controller (SC) for counting statistics, and
 - 25 a port communications controller (PCC) for enabling port controllers to communicate.
8. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 7, the address validation/translation controller further comprising:
- 30 a timer for timing and timeout handling of frames in the queue, such that upon passing a buffer descriptor to the destination F_port, the port controller from the source F_port remembers the frame.
9. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 8, wherein the port controller, in response to the timer, sends a busy frame signal back to the source N_Port, or discards the frame, if the frame cannot be forwarded out to the destination F_port within a fixed time period.
10. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 9, wherein the fixed time period is determined in accordance with the Fibre channel standard.
11. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 10, wherein the source F_port retains responsibility for timing the frame until the source F_port buffer descriptor makes it to the head of the source queue and is passed to the destination queue, such that the destination port may now forward a request to the source F_port to return a busy frame signal.
12. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 10, further comprising a "non-core function port" port controller, wherein, the buffer descriptor of the frame to be busied or rejected is queued to the "non-core function port" port controller, wherein the source F_port is not required to generate busy and reject frame commands.
13. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 2, the plurality of port controllers further comprising a "non-core function port" port controller, wherein, the buffer descriptor of the frame to be busied or rejected is queued to the "non-core function port" port controller, wherein the source F_port is not required to generate busy and reject frame commands.
14. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 13, the "non-core function port" port controller further comprising software for manipulation of the frame header, and further comprising the step of queuing the frame back out to the original source F_port, and wherein, the source F_port continues to time the frame while the original frame is queued to the "non-core function port", and wherein, if the frame is not forwarded within a pre-determined period of time, the source F_port makes a request to the "non-core function port" port controller not to forward the original frame to the processor, and wherein, if the frame does make it the processor, and the reject/busy frame is queued back to the source F_port, the "non-core function port" port controller must tie the frame and discard it if necessary.
15. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 14, further comprising a "non-core function port" timer that times the time between when the frame was first queued to the "non-core function" port and when it was actually delivered to it.
16. The fibre channel switch employing a distributed source and destination queuing algorithm as claimed in claim 5, further comprising a processor and the buffer information passed to the "non-core function port" further comprises a time stamp, wherein, the processor time stamps frames upon receipt, and wherein, a time out value is calculated that ensures that the frame is discarded within R_A_TOV.

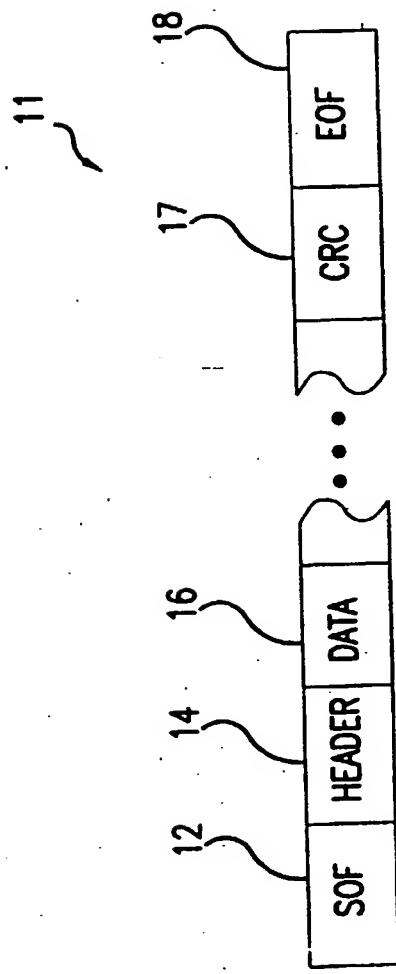


FIG. 1

PRIOR ART

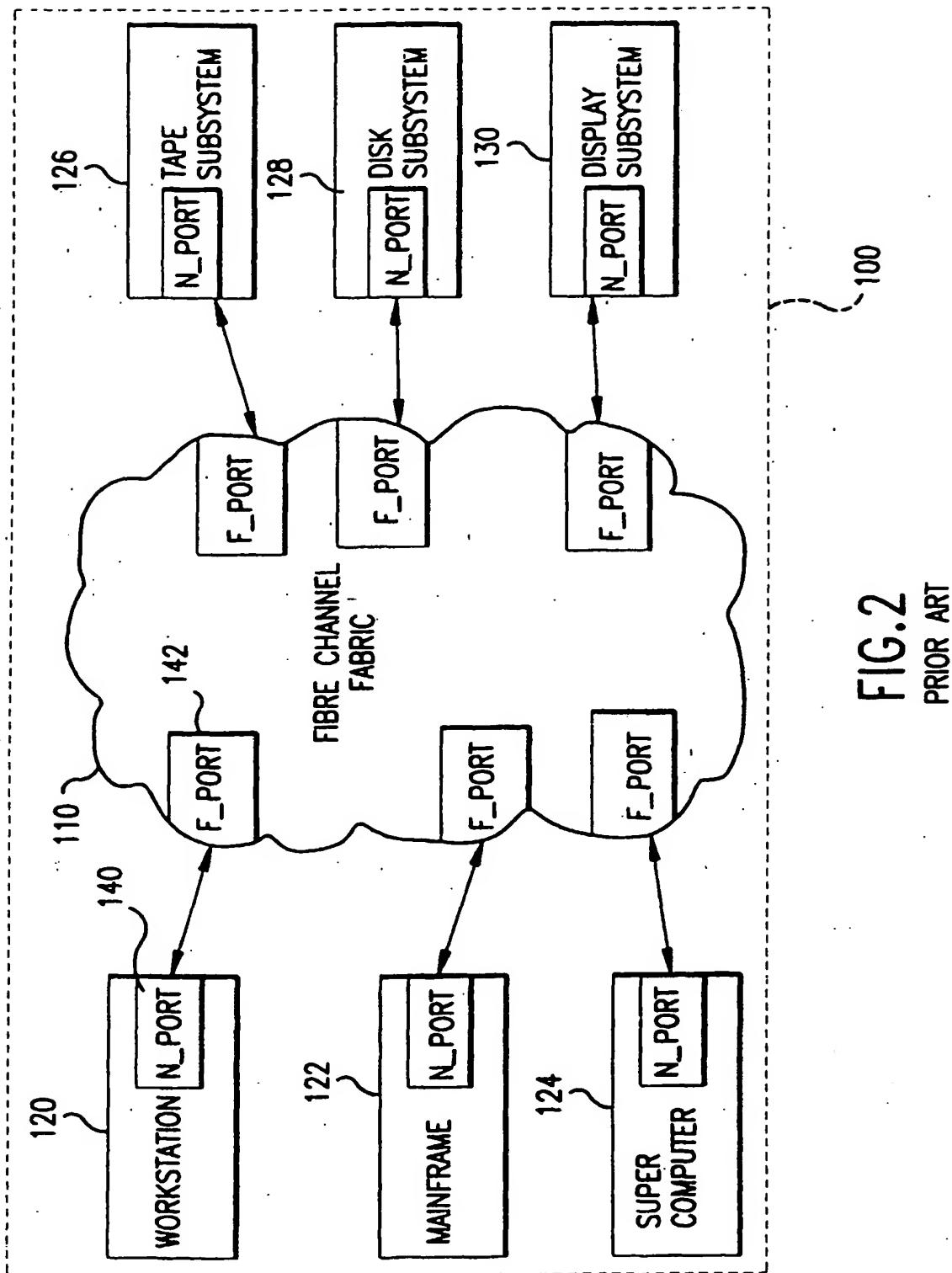
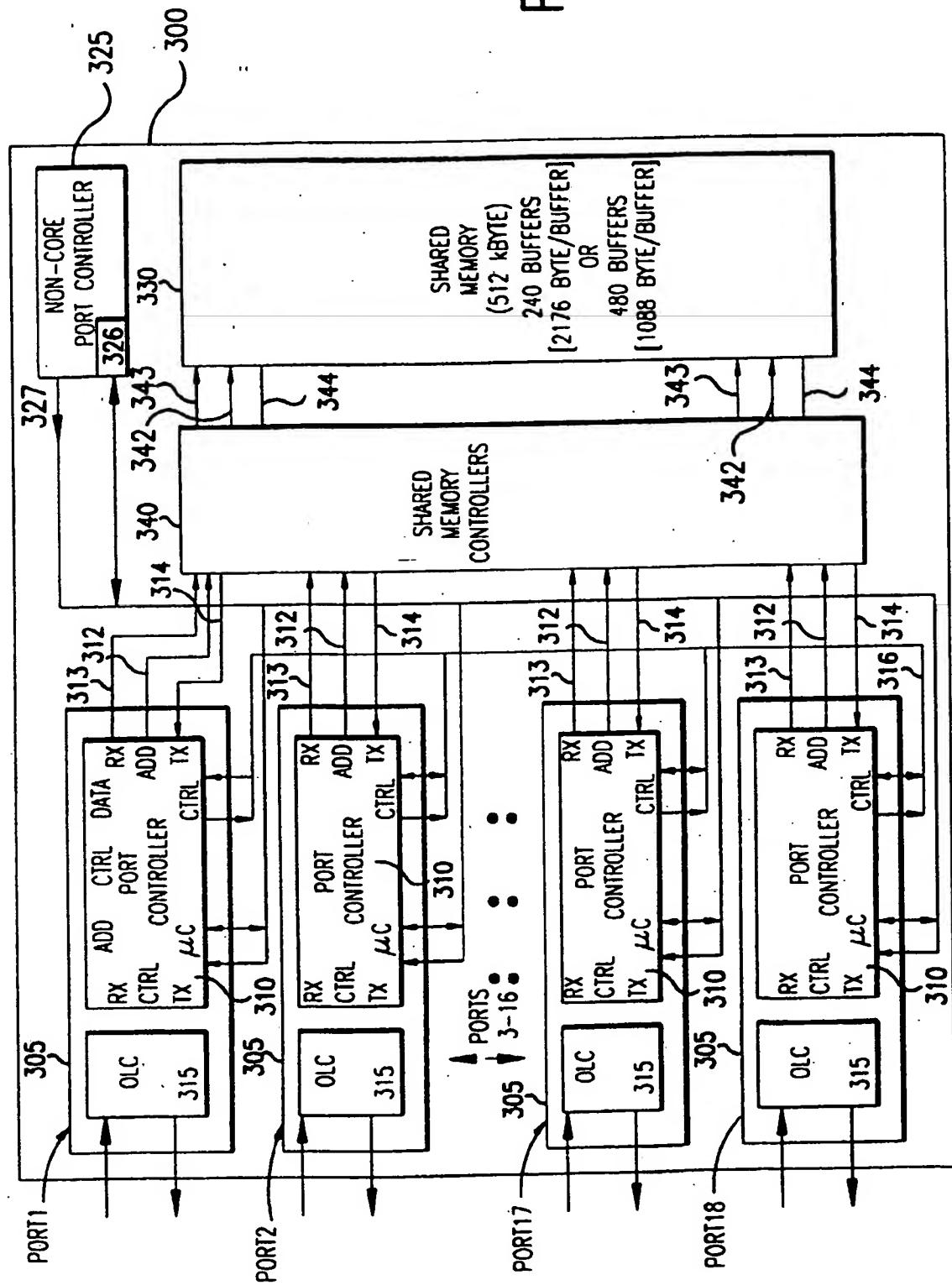


FIG.2
PRIOR ART

FIG. 3



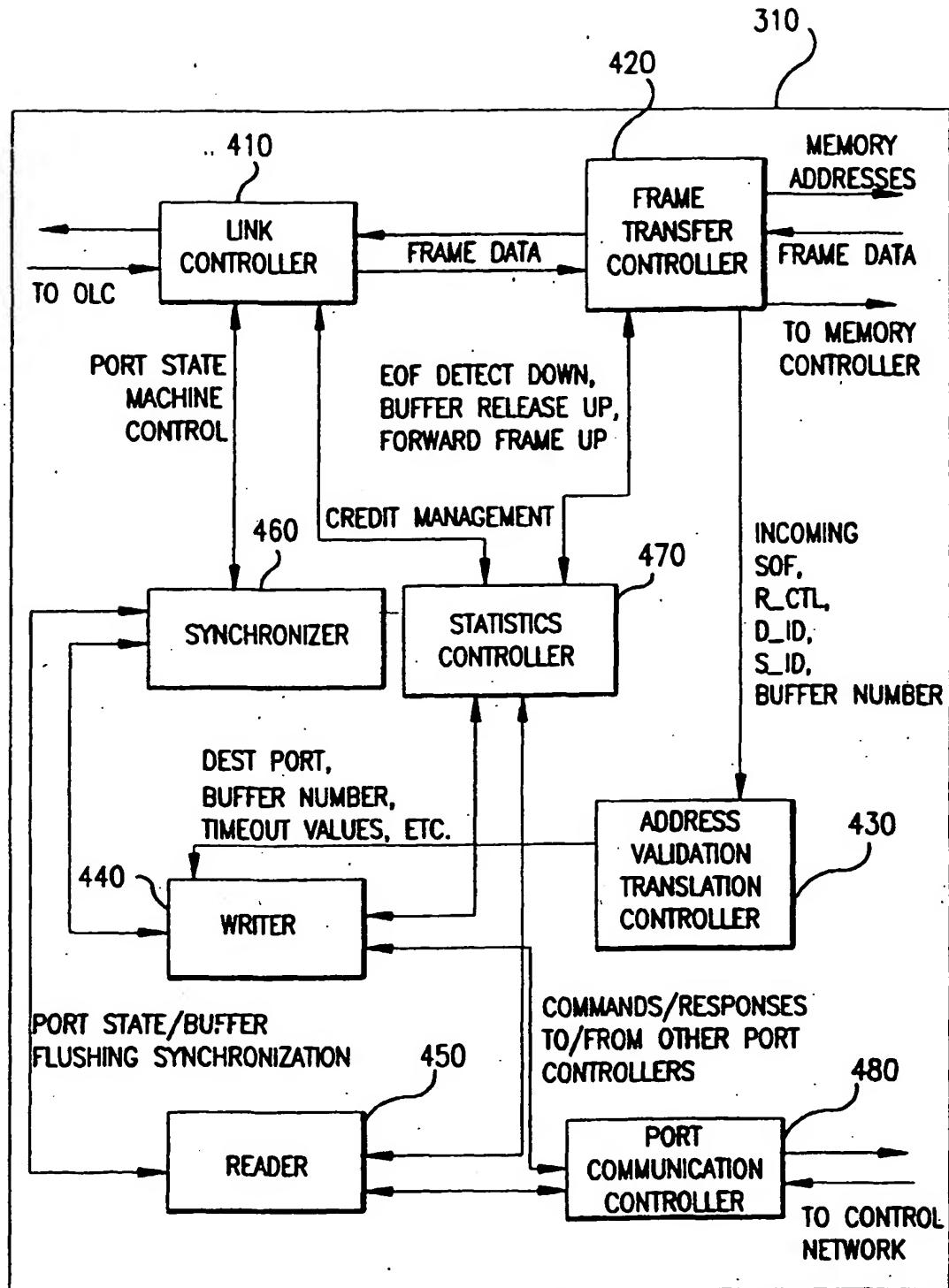
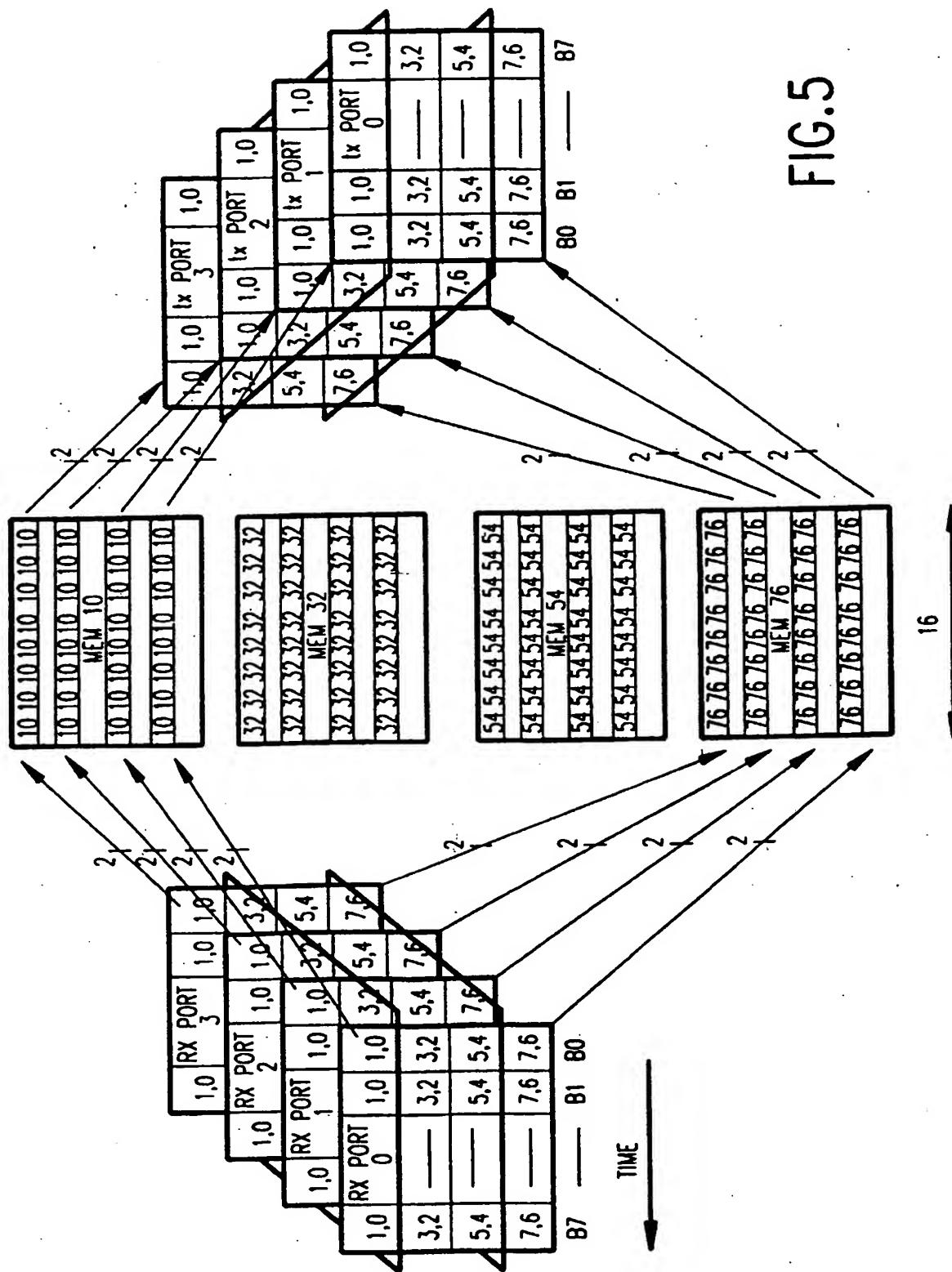


FIG.4

FIG. 5



(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets

(11)



EP 0 841 781 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
11.09.2002 Bulletin 2002/37

(51) Int Cl.7: H04L 12/44

(43) Date of publication A2:
13.05.1998 Bulletin 1998/20

(21) Application number: 97630053.3

(22) Date of filing: 05.09.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

Designated Extension States:
AL LT LV RO SI

(30) Priority: 11.09.1996 US 714029

(71) Applicant: McData Corporation
Broomfield, Colorado 80021-3464 (US)

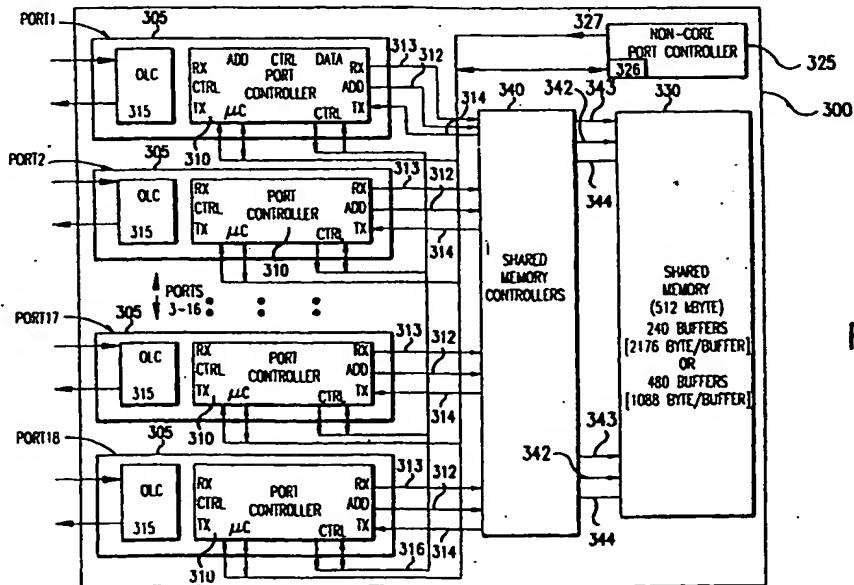
(72) Inventor: Book, David
Thornhill, Ontario L4J 7X1 (CA)

(74) Representative: Driver, Virginia Rozanne
Page White & Farrer
54 Doughty Street
London WC1N 2LS (GB)

(54) Fibre channel switch employing distributed queuing

(57) The present invention is a fibre channel switch employing a distributed queuing algorithm for interconnecting a plurality of devices (workstations, supercomputer, peripherals) through their associated node ports (N_ports) and employs a fabric having a shared memory coupled to a plurality of fabric ports (F_ports) through a bi-directional bus over which memory addresses, frame data and communications commands are transmitted.

Each F_port includes a port controller employing a distributed queuing algorithm associated with a control network for communicating commands between the ports related to when and where frame transfers should be made, wherein the bi-directional bus provides an independent data network for access to the shared memory such that frames can be transferred to and from the shared memory in response to port controller commands.



EP 0 841 781 A3



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 97 63 0053

DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim
D, X	EP 0 709 986 A (HEWLETT-PACKARD COMPANY) 1 May 1996 (1996-05-01)	1-3, 6
A	* column 8, line 33 - column 9, line 23 * * column 10, line 8 - line 23 * * column 11, line 16 - line 19 * * column 17, line 5 - line 45 * * figure 9 *	7, 12-16
A	BIRMAN A: "A HIGH-PERFORMANCE SWITCH WITH APPLICATIONS TO FRAME RELAY NETWORKS" MULTIPLE FACETS OF INTEGRATION. SAN FRANCISCO, JUNE 3 - 7, 1990, PROCEEDINGS OF THE ANNUAL JOINT CONFERENCE OF THE COMPUTER AND COMMUNICATIONS SOCIETIES (INFOCOM), WASHINGTON, IEEE COMP. SOC. PRESS, US, vol. 2 CONF. 9, 3 June 1990 (1990-06-03), pages 782-789, XP000164297 ISBN: 0-8186-2049-8 * page 783, left-hand column, line 9 - line 19 *	4, 5
A	DENZEL W E ET AL: "A HIGHLY MODULAR PACKET SWITCH FOR GB/S RATES" PROCEEDINGS OF THE INTERNATIONAL SWITCHING SYMPOSIUM. YOKOHAMA, OCT. 25 - 30, 1992, TOKYO, IEICE, JP, vol. 2 SYMP. 14, 25 October 1992 (1992-10-25), pages 236-240, XP000337725 * sections 2.2.1 and 2.2.2 * * figure 2 *	4, 5
A	US 5 455 825 A (GHOSH ABHIJIT ET AL) 3 October 1995 (1995-10-03) * column 13, line 38 - column 14, line 32 * * claim 1 *	8-11
The present search report has been drawn up for all claims		
Place of search	Date of completion of the search	Examiner
THE HAGUE	28 June 2002	Perez Perez, J
CATEGORY OF CITED DOCUMENTS		
X : particularly relevant if taken alone	T : theory or principle underlying the invention	
Y : particularly relevant if combined with another document of the same category	E : earlier patent document, but published on, or after the filing date	
A : technological background	D : document cited in the application	
G : non-written disclosure	L : document cited for other reasons	
P : intermediate document	& : member of the same patent family, corresponding document	

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 97 63 0053

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

28-06-2002

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
EP 709986	A	01-05-1996	US CA EP JP	5603064 A 2153823 A1 0709986 A2 8274720 A	11-02-1997 28-04-1996 01-05-1996 18-10-1996
US 5455825	A	03-10-1995	JP	8056230 A	27-02-1996

EPO FORM P059

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

THIS PAGE BLANK (USPTO)